Announcements

Monday, November 20

- You already have your midterms!
  - Course grades will be curved at the end of the semester. The percentage of A's, B's, and C's to be awarded depends on many factors, and will not be determined until all grades are in.
  - Individual exam grades are not curved.

- WeBWorK 6.1, 6.2, 6.3 due the Wednesday after Thanksgiving.

- **Reading day:** Math is 1–3pm on December 6 in Clough 144 and 152. I'll be there for part of it.

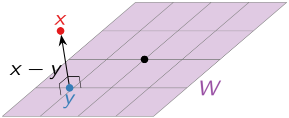
- My office is Skiles 244. Rabinoffice hours are Monday, 1–3pm and Tuesday, 9–11am.

Section 6.2/6.3

Orthogonal Projections

Best Approximation

Suppose you measure a data point x which you know for theoretical reasons must lie on a subspace W.



Due to measurement error, though, the measured x is not actually in W. Best approximation: y is the *closest* point to x on W.

How do you know that y is the closest point? The vector from y to x is orthogonal to W: it is in the *orthogonal complement* $W^{\perp}$.

Orthogonal Decomposition

Recall: If W is a subspace of $\mathbf{R}^n$, its **orthogonal complement** is

$$W^\perp = \{v \text{ in } \mathbf{R}^n \mid v \text{ is perpendicular to every vector in } W\}$$

Theorem

Every vector x in $\mathbf{R}^n$ can be written as

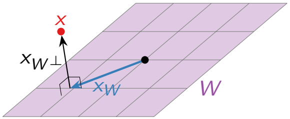
$$x = x_W + x_{W^\perp}$$

for unique vectors x_W in W and $x_{W^\perp}$ in $W^\perp$.

The equation $x = x_W + x_{W^\perp}$ is called the **orthogonal decomposition** of x (with respect to W).

The vector x_W *is the closest vector to* x *on* W.

[interactive 1] [interactive 2]



Theorem

Every vector x in $\mathbf{R}^n$ can be written as

$$x = x_W + x_{W^\perp}$$

for unique vectors x_W in W and $x_{W^\perp}$ in $W^\perp$.

Why?

Uniqueness: suppose $x = x_W + x_{W^\perp} = x'_W + x'_{W^\perp}$ for x_W, x'_W in W and $x_{W^\perp}, x'_{W^\perp}$ in $W^\perp$. Rewrite:

$$x_W - x'_W = x'_{W^\perp} - x_{W^\perp}.$$

The left side is in W, and the right side is in $W^\perp$, so they are both in $W \cap W^\perp$. But the only vector that is perpendicular to itself is the zero vector! Hence

$$0 = x_W - x'_W \implies x_W = x'_W$$
$$0 = x_{W^\perp} - x'_{W^\perp} \implies x_{W^\perp} = x'_{W^\perp}$$

Existence: We will compute the orthogonal decomposition later using orthogonal projections.

Orthogonal Decomposition

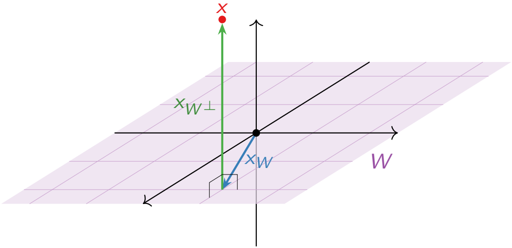
Example

Let W be the xy-plane in $\mathbf{R}^3$. Then $W^\perp$ is the z-axis.

$$x = \begin{pmatrix}1\\2\\3\end{pmatrix} \implies x_W = \begin{pmatrix}1\\2\\0\end{pmatrix} \qquad x_{W^\perp} = \begin{pmatrix}0\\0\\3\end{pmatrix}.$$

$$x = \begin{pmatrix}a\\b\\c\end{pmatrix} \implies x_W = \begin{pmatrix}a\\b\\0\end{pmatrix} \qquad x_{W^\perp} = \begin{pmatrix}0\\0\\c\end{pmatrix}.$$

This is just decomposing a vector into a "horizontal" component (in the xy-plane) and a "vertical" component (on the z-axis).



[interactive]

Orthogonal Decomposition

Computation?

Problem: Given x and W, how do you compute the decomposition $x = x_W + x_{W^\perp}$?

Observation: It is enough to compute x_W, because $x_{W^\perp} = x - x_W$.

First we need to discuss orthogonal sets.

Definition

A set of *nonzero* vectors is **orthogonal** if each pair of vectors is orthogonal. It is **orthonormal** if, in addition, each vector is a unit vector.

Lemma

An orthogonal set of vectors is linearly independent. Hence it is a basis for its span.

Suppose $\{u_1, u_2, \ldots, u_m\}$ is orthogonal. We need to show that the equation

$$c_1 u_1 + c_2 u_2 + \cdots + c_m u_m = 0$$

has only the trivial solution $c_1 = c_2 = \cdots = c_m = 0$.

$$0 = u_1 \cdot (c_1 u_1 + c_2 u_2 + \cdots + c_m u_m) = c_1(u_1 \cdot u_1) + 0 + 0 + \cdots + 0.$$

Hence $c_1 = 0$. Similarly for the other c_i.

Example: $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$ is an orthogonal set. Check:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0 \qquad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0 \qquad \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0.$$

Example: $\mathcal{B} = \{e_1, e_2, e_3\}$ is an orthogonal set. Check:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0 \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0.$$

Example: Let $x = \binom{a}{b}$ be a nonzero vector, and let $y = \binom{-b}{a}$. Then $\{x, y\}$ is an orthogonal set:

$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} -b \\ a \end{pmatrix} = -ab + ab = 0.$$

Orthogonal Projections

Definition

Let W be a subspace of $\mathbf{R}^n$, and let $\{u_1, u_2, \ldots, u_m\}$ be an *orthogonal* basis for W. The **orthogonal projection** of a vector x onto W is

$$\operatorname{proj}_W(x) \stackrel{\text{def}}{=} \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i = \frac{x \cdot u_1}{u_1 \cdot u_1}\, u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2}\, u_2 + \cdots + \frac{x \cdot u_n}{u_n \cdot u_n}\, u_n.$$

This is a vector in W because it is in $\operatorname{Span}\{u_1, u_2, \ldots, u_m\}$.

Theorem

Let W be a subspace of $\mathbf{R}^n$, and let x be a vector in $\mathbf{R}^n$. Then

$$x_W = \operatorname{proj}_W(x) \quad \text{and} \quad x_{W^\perp} = x - \operatorname{proj}_W(x).$$

In particular, $\operatorname{proj}_W(x)$ is the closest point to x in W.

Why? Let $y = \operatorname{proj}_W(x)$. We need to show that $x - y$ is in $W^\perp$. In other words, $u_i \cdot (x - y) = 0$ for each i. Let's do u_1:

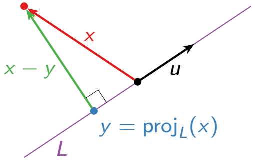
$$u_1 \cdot (x - y) = u_1 \cdot \left(x - \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i \right) = u_1 \cdot x - \frac{x \cdot u_1}{u_1 \cdot u_1}(u_1 \cdot u_1) - 0 - \cdots = 0.$$

Orthogonal Projection onto a Line

The formula for orthogonal projections is simple when W is a *line*.

Let $L = \text{Span}\{u\}$ be a line in $\mathbf{R}^n$, and let x be in $\mathbf{R}^n$. The orthogonal projection of x onto L is the point

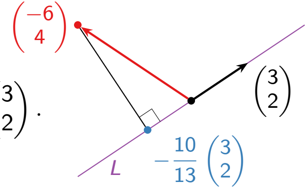
$$\text{proj}_L(x) = \frac{x \cdot u}{u \cdot u}\, u.$$



[interactive]

Example: Compute the orthogonal projection of $x = \binom{-6}{4}$ onto the line L spanned by $u = \binom{3}{2}$.

$$y = \text{proj}_L(x) = \frac{x \cdot u}{u \cdot u} u = \frac{-18+8}{9+4}\begin{pmatrix}3\\2\end{pmatrix} = -\frac{10}{13}\begin{pmatrix}3\\2\end{pmatrix}.$$



Orthogonal Projection onto a Plane

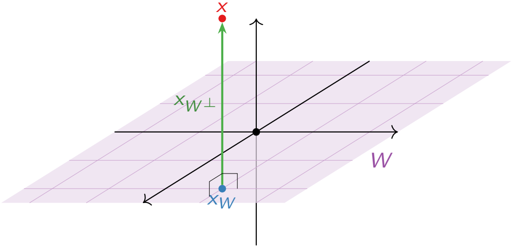
Easy example

What is the projection of $x = \begin{pmatrix}1\\2\\3\end{pmatrix}$ onto the xy-plane?

Answer: The xy-plane is $W = \mathrm{Span}\{e_1, e_2\}$, and $\{e_1, e_2\}$ is an orthogonal basis.

$$x_W = \mathrm{proj}_W \begin{pmatrix}1\\2\\3\end{pmatrix} = \frac{x \cdot e_1}{e_1 \cdot e_1} e_1 + \frac{x \cdot e_2}{e_2 \cdot e_2} e_2 = \frac{1 \cdot 1}{1^2} e_1 + \frac{1 \cdot 2}{1^2} e_2 = \begin{pmatrix}1\\2\\0\end{pmatrix}.$$

So this is the same projection as before.



[interactive]

Orthogonal Projections

More complicated example

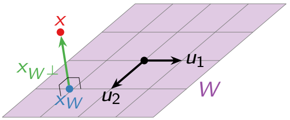
What is the projection of $x = \begin{pmatrix} -1.1 \\ 1.4 \\ 1.45 \end{pmatrix}$ onto $W = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1.1 \\ -.2 \end{pmatrix} \right\}$?

Answer: The basis is orthogonal, so

$$x_W = \text{proj}_W \begin{pmatrix} -1.1 \\ 1.4 \\ 1.45 \end{pmatrix} = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2$$

$$= \frac{(-1.1)(1)}{1^2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \frac{(1.4)(1.1) + (1.45)(-.2)}{1.1^2 + (-.2)^2} \begin{pmatrix} 0 \\ 1.1 \\ -.2 \end{pmatrix}$$

This turns out to be equal to $u_2 - 1.1u_1$.



[interactive]

First we restate the property we've been using all along.

Let x be a vector and let $x = x_W + x_{W^\perp}$ be its orthogonal decomposition with respect to a subspace W. The following vectors are the same:

- x_W
- $\text{proj}_W(x)$
- The closest vector to x on W

We can think of orthogonal projection as a *transformation*:

$$\text{proj}_W : \mathbf{R}^n \longrightarrow \mathbf{R}^n \qquad x \mapsto \text{proj}_W(x).$$

Theorem

Let W be a subspace of $\mathbf{R}^n$.

1. proj_W is a *linear* transformation.
2. For every x in W, we have $\text{proj}_W(x) = x$.
3. For every x in $W^\perp$, we have $\text{proj}_W(x) = 0$.
4. The range of proj_W is W and the null space of proj_W is $W^\perp$.

Poll

Let W be a subspace of $\mathbf{R}^n$.

Poll

Let A be the matrix for proj_W. What is/are the eigenvalue(s) of A?

A. 0 B. 1 C. -1 D. $0, 1$ E. $1, -1$ F. $0, -1$ G. $-1, 0, 1$

The 1-eigenspace is W.

The 0-eigenspace is $W^\perp$.

We have $\dim W + \dim W^\perp = n$, so that gives n linearly independent eigenvectors already.

So the answer is D.

What is the matrix for $\mathrm{proj}_W \colon \mathbf{R}^3 \to \mathbf{R}^3$, where

$$W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}?$$

Answer: Recall how to compute the matrix for a linear transformation:

$$A = \begin{pmatrix} | & | & | \\ \mathrm{proj}_W(e_1) & \mathrm{proj}_W(e_2) & \mathrm{proj}_W(e_3) \\ | & | & | \end{pmatrix}.$$

We compute:

$$\mathrm{proj}_W(e_1) = \frac{e_1 \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{e_1 \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5/6 \\ 1/3 \\ -1/6 \end{pmatrix}$$

$$\mathrm{proj}_W(e_2) = \frac{e_2 \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{e_2 \cdot u_2}{u_2 \cdot u_2} u_2 = 0 + \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

$$\mathrm{proj}_W(e_3) = \frac{e_3 \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{e_3 \cdot u_2}{u_2 \cdot u_2} u_2 = -\frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/6 \\ 1/3 \\ 5/6 \end{pmatrix}$$

Therefore $A = \begin{pmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{pmatrix}$.

Orthogonal Projections

Matrix facts

Let W be an m-dimensional subspace of $\mathbf{R}^n$, let $\text{proj}_W : \mathbf{R}^n \to W$ be the projection, and let A be the matrix for proj_L.

Fact 1: A is diagonalizable with eigenvalues 0 and 1; it is similar to the diagonal matrix with m ones and $n - m$ zeros on the diagonal.

Why? Let $v_1, v_2, \ldots, v_m$ be a basis for W, and let $v_{m+1}, v_{m+2}, \ldots, v_n$ be a basis for $W^\perp$. These are (linearly independent) eigenvectors with eigenvalues 1 and 0, respectively, and they form a basis for $\mathbf{R}^n$ because there are n of them.

Example: If W is a plane in $\mathbf{R}^3$, then A is similar to projection onto the xy-plane:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Fact 2: $A^2 = A$.

Why? Projecting twice is the same as projecting once:

$$\text{proj}_W \circ \text{proj}_W = \text{proj}_W \implies A \cdot A = A.$$

Coordinates with respect to Orthogonal Bases

Let W be a subspace with orthogonal basis $\mathcal{B} = \{u_1, u_2, \ldots, u_m\}$.

For x in W we have $\text{proj}_W(x) = x$, so

$$x = \text{proj}_W(x) = \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i = \frac{x \cdot u_1}{u_1 \cdot u_1}\, u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2}\, u_2 + \cdots + \frac{x \cdot u_n}{u_n \cdot u_n}\, u_n.$$

This makes it easy to compute the $\mathcal{B}$-coordinates of x.

Corollary

Let W be a subspace with orthogonal basis $\mathcal{B} = \{u_1, u_2, \ldots, u_m\}$. Then

$$[x]_{\mathcal{B}} = \left(\frac{x \cdot u_1}{u_1 \cdot u_1},\ \frac{x \cdot u_2}{u_2 \cdot u_2},\ \ldots,\ \frac{x \cdot u_m}{u_m \cdot u_m} \right).$$

[interactive]

Problem: Find the $\mathcal{B}$-coordinates of $x = \binom{0}{3}$, where

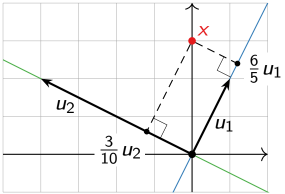
$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \end{pmatrix} \right\}.$$

Old way:

$$\left(\begin{array}{cc|c} 1 & -4 & 0 \\ 2 & 2 & 3 \end{array}\right) \xrightarrow{\text{rref}} \left(\begin{array}{cc|c} 1 & 0 & 6/5 \\ 0 & 1 & 6/20 \end{array}\right) \implies [x]_{\mathcal{B}} = \begin{pmatrix} 6/5 \\ 6/20 \end{pmatrix}.$$

New way: note $\mathcal{B}$ is an *orthogonal* basis.

$$[x]_{\mathcal{B}} = \left(\frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 \right) = \left(\frac{3 \cdot 2}{1^2 + 2^2}, \frac{3 \cdot 2}{(-4)^2 + 2^2} \right) = \left(\frac{6}{5}, \frac{3}{10} \right).$$



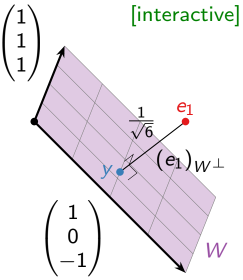
[interactive]

What is the distance from e_1 to $W = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$?

Answer: The closest point on W to e_1 is $\text{proj}_W(e_1) = \begin{pmatrix} 5/6 \\ 1/3 \\ -1/6 \end{pmatrix}$.

The distance from e_1 to this point is

$$\begin{aligned}
\text{dist}\big(e_1, \text{proj}_W(e_1)\big) &= \|(e_1)_{W^\perp}\| \\
&= \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 5/6 \\ 1/3 \\ -1/6 \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} 1/6 \\ -1/3 \\ -1/6 \end{pmatrix} \right\| \\
&= \sqrt{(1/6)^2 + (-1/3)^2 + (-1/6)^2} \\
&= \frac{1}{\sqrt{6}}.
\end{aligned}$$



Summary

Let W be a subspace of $\mathbf{R}^n$.

- Any vector x in $\mathbf{R}^n$ can be written in a unique way as

$$x = x_W + x_{W^\perp}$$

  for x_W in W and $x_{W^\perp}$ in $W^\perp$. This is called its **orthogonal decomposition.**
- The vector x_W is the *closest point to x in W*: it is the *best approximation*.
- The *distance* from x to W is $\|x_{W^\perp}\|$.
- If you have an *orthogonal* basis $\{u_1, u_2, \ldots, u_m\}$ for W, then

$$x_W = \text{proj}_W(x) = \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i = \frac{x \cdot u_1}{u_1 \cdot u_1}\, u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2}\, u_2 + \cdots + \frac{x \cdot u_n}{u_n \cdot u_n}\, u_n.$$

  Hence $x_{W^\perp} = x - \text{proj}_W(x)$.
- If you have an *orthogonal* basis $\{u_1, u_2, \ldots, u_m\}$ for W, then

$$[x]_{\mathcal{B}} = \left(\frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \ldots, \frac{x \cdot u_m}{u_m \cdot u_m} \right).$$

- We can think of $\text{proj}_W \colon \mathbf{R}^n \to \mathbf{R}^n$ as a linear transformation. Its null space is $W^\perp$, and its range is W.
- The matrix A for proj_W is diagonalizable with eigenvalues 0 and 1. It is *idempotent*: $A^2 = A$.